Applied Soft Computing Journal 99 (2021) 106898

Contents lists available at ScienceDirect

Applied Soft Computing Journal

journal homepage: www.elsevier.com/locate/asoc

Review article

Financial market prediction under deep learning framework using auto encoder and kernel extreme learning machine

D.K. Mohanty, Ajaya Kumar Parida *, Shelly Suman Khuntia

School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha, India

A R T I C L E I N F O

Article history:
Received 29 June 2020
Received in revised form 3 November 2020
Accepted 5 November 2020
Available online 10 November 2020

Keywords:
Prediction methods
Stacked AE
KELM
HELM and stock market

A B S T R A C T

The technical indicators are highly uncertain therefore possess greater influence on the stock market prediction. Among different techniques developed for effective prediction of the financial market the AI techniques show better prediction efficiency. In this paper, a hybrid model combined with auto encoder (AE) and kernel extreme learning machine (KELM) is proposed for further improvement in the quality of financial market prediction. This study mainly emphasizes on a precise prediction of the financial market, the main motive behind stock price prediction is minimizing the substantial losses faced by investors, and analysing the profitability with the help of buying and selling amount. The prime advantage of the proposed technique over the conventional SAE is robust prediction of different financial market with reduction in error. To authenticate the performance of the proposed deep learning (DL) technique (KELM-AE), high-frequency data of different financial market like Yes Bank, SBI, ASHR, and DJI are taken into consideration and the performance of the proposed technique is investigated in MATLAB based simulation in accordance with MAPE (Mean Absolute Percentage Error), MAE (Mean Absolute Error) and RMSE (Root Mean Square Error). The application of SAE is new in the field of predicting different bank data. The validation of the model is performed by comparing it with other traditional methods based on different performance indexes. The simulation result indicates that the proposed DL based technique (KELM-AE) outperforms other models with a MAPE value of less than 2%for future prediction, irrespective of the financial market. For example the MAPE value for KELM-AE is observed to be 1.074 %, 0.888%, 1.021% for YES, SBI and BOI respectively which is much lower as compared to other model like ELM that shows a MAPE value of 1.714%, 1.473% and 1.550% for the above mentioned bank.

Contents

* Corresponding author.
E-mail addresses: dkmohanty.iitkgp@gmail.com (D.K. Mohanty), erakparida@gmail.com (A.K. Parida), shelly.khuntia@gmail.com (S.S. Khuntia).

https://doi.org/10.1016/j.asoc.2020.106898
1568-4946/© 2020 Published by Elsevier B.V.

Abbreviation

AE	Auto Encoder
ARMA	Auto-Regressive Moving Average
ARIMA	Auto-Regressive Integrated Moving Average
ARMAX	Auto-Regressive Moving Average with Exogenous terms
ANN	Artificial Neural Network
AI	Artificial Intelligence
BP	Back Propagation
CNN	Convolutional Neural Network
DL	Deep Learning
ELM	Extreme Learning Machine
FL	Fuzzy Logic
FFNN	Feed Forward Neural Network
fmp	financial market price feed as input to the model
GARCH	Generalized Auto-Regressive Conditional Heteroskedasticity
HELM	Hierarchical Extreme Learning Machine
KELM	Kernel Extreme Learning Machine
KKT	Karush–Kuhn–Tucker theorem
LSTM	Long Short Term Memory
MAPE	Mean Absolute Percentage Error
MAE	Mean Absolute Error
NSE	National Stock Exchange
NYSE	New York Stock Exchange
OH LC	Open High Low Close
RKELM	Robust Kernel Extreme Learning Machine
RMSE	Root Mean Square Error
RBF	Radial Basis Functional neural n network
SLFN	Single Hidden layer Feed Forward Neural Network
SAE	Stacked Auto encoder
PO	Predicted Output
AO	Actual Output

List of Symbols

Symbols	**Full form**
α	Lagrange multiplier
ξ	Error
β	Output weight vector
b_j	Bias values of the hidden layer
ε	Non zero training error
σ, bc, d, e, f and γ	Kernel functions
a	Activation function
C	Regularization parameter
d_t	Present actual value
d_{t+1}	Predicted value of the next time period.
h_i	Feature mapping
H	Hidden layer output matrix
H^+	pseudo-inverse matrix of H
I	Unit matrix
N_{norm}	Normalized data
N_{min}	Minimum price of the data
N_{max}	Maximum price of the data
N	Total number of samples
O_i	Measured output at ith instant
o_p	Output matrix
t_i	Target at ith instant
T	Target output
w_j	Input weight
W_1	Weight matrix of the hidden layer
W_2	Weight matrix of the reconstruction layer
x_i	ith input
X	Total data set
R_s	Strategy returns
s	Is the total number of days for selling
b	Is the total number of days for buying
S_c	Transaction cost for selling
B_c	Transaction cost for buying

1. Introduction

Stock market and its prediction are treated as one of the classic topic for both the financial market as well as the academic circle. The stock price is greatly affected by different factors like global economic status, political factors, and some other fundamental factors hence a constructive approach for predicting the highly fluctuating financial market is of utmost importance. Precise prediction of the financial market is beneficial for various financial decision making purposes therefore a well-founded prediction method is needed for prediction of market fluctuations. Previous literature studies showed various controversies related to the prediction of stock market. In the previous studies random walk theory was implemented for determining the stock price movement which was further contradicted with the EMH method that proposed the impossibility of stock market prediction using the past data. Further different studies showed the prediction of stock price is possible with the help of previous data; numerous predictive attempts made in the literature survey for precise price prediction are briefly described in this section. Previously various statistical techniques were implemented for the purpose of price prediction, some of such models are ARMA [1], where the changes in the volatility of S&P 500 market is observed using the ARMA model. ARIMA [2], in this the authors emphasizes on the usefulness of forecasting mid-term price trend of Taiwan stock market. The basic of the system is recurrent neural network trained with the help of different features extracted by analysing the ARIMA model. GARCH [3,4], technique were also implemented for performing various prediction in different field of application. The main limitation of this technique is its incapability of better performance when nonlinear data is taken into consideration as the stock market data is highly noisy in nature; these drawbacks were further minimized with the help of different AI techniques such as ANN, SVM, FL [5–8], among which the ANN is the mostly used method because of its capability to solve highly nonlinear data of various field. The literature survey implies the use of different AI techniques for accurate stock market prediction, techniques like ANN, SVM [9,10], found a wide area of application and also showed better result when compared to the very basic models and in [11] predicted values of GEP and SVM methods were implemented as two new potential and novel approaches using R and RMSE for comparing results. RBFLN, FFNN, fuzzy logic [5,6], ELM etc. Showed better performance when nonlinear data are considered. ANN has a wide area of implementation in the field of prediction [12–18], further different hybrid techniques were introduced combining the ANN methods like fuzzy ANN, GARCH-ANN [19–25]. Another introduction observed in the previous literature study is the FFNN, which has been implemented in different field of study [24,25]. SLFN is the most popular feed forward technique that uses different training algorithm for both prediction and classification purposes [26,27]. A new machine learning technique known as the extreme learning machine (ELM) came into existence for solving the problem of local minima and low convergence rate [28–30]. It is a non iterative SLFN process with less computational time but high prediction accuracy. ELM basically depends on propitious selection of the hidden neurons and the activation function for the purpose of stability and generalization. This problem leads to another research contribution with the absence of feature mapping functions in ELM and thus kernel functions came into existence in combination with the ELM, and can be used for better achievements with improvement in stability and generalization. The primary notion of the KELM technique is given in ref [31–35], where it shows that there is no need of choosing the random weight thus leading to more accuracy in price prediction with comparatively less computational time. The main disadvantage of this technique is its choice of the kernel function and the corresponding values assigned to it. Lately the DL techniques have gained more attention in the area of image processing, speech recognition classification etc., they even performs better than most other AI techniques when a variety of inputs are considered [36,37].

This study emphasizes on developing a model that gives a more accurate prediction of the financial market with less computational complexity. The DL based model is implemented in this paper to obtain a more accurate price prediction. DL is a multi neural network consisting of cascaded layers of no linear processing unit from which the features are extracted. It can be initialized by both supervised and unsupervised learning [36,38]. The sophisticated algorithm of the DL technique not only improves the computational capability but also gives the power to solve big and complicated data. Multiple AE are stacked together and termed as the stacked AE (SAE) [39], SAE is implemented to achieve the goal of higher accuracy in price prediction. The ELM technique is widely incorporated with the AE for prediction purpose [39]. In case of the financial market, they extract abstract features without any interference of control variables [40–43]. Various studies showed the performance superiority of the DL technique [44–48]. Different DL architectures have been discussed in the later section of this study. Ref [49] gives a wide knowledge of various prediction techniques implemented in the literature and the better performing among them. The authors have emphasized on different techniques for predicting financial market and also proved that the DL methods performed way better than their counter machine learning parts, the authors have also showed the increasing interest in the concerned technique in all areas of research.

This paper proposes the application of AE for layer wise training of different input variables like OH LC [50,51]. This study deals with SAE for prediction purpose where the last layer is formed of KELM method for decision making purpose. Irrespective of all the advantages the main disadvantage of the AE method is that along with reduction in the data dimension it also causes loss of data. In [52] the authors have derived a novel technique for the prediction of currency but the main issue that is of utmost concern is the EWT based decomposition that leads to larger run time and also a matter of concern when larger data set is taken into consideration, while in this proposed model the multilayer network of the SAE removes the noise present in the signal and can perform significantly. In future studies a much more advanced model can be combined along with the DL technique for optimizing various kernel parameters when various market fluctuation might also be considered for example fluctuation of financial market due to COVID situation [53].

Prediction of stock market is essential as there are various factors like economic fluctuations, political issues etc. that introduces various uncertainties to the model, a thorough analysis helps the investors in decision making process to be at low risk for losses by determining the upward or downward movement of the stock price. One of the important matters in the stock market is its predicted price and estimated real value of different companies listed, since prices are a signal to guide the effective tool of the financial market, leading to better profit for both the speculator and the investor.

Following the previous literature survey the key contribution of this study adding newness to the previously implemented technique is briefed as follows:

- Introduction of a modified multilayer-ELM, where the encoding is performed with the help of randomized AE, here the multiple layers are stacked but the final classifier is the Kernel based ELM. The use of SAE for financial market prediction has not been previously implemented.
- A precise prediction of stock market using this AE-KELM is performed with comprehensive experiments based on different financial data.
- Prediction is made with four input variables commonly known as the OHLC
- Movement of the stock price on daily basis is performed for minimizing substantial losses faced by investors.
- Promising prediction accuracy is observed for the proposed technique when compared with different other previously implemented techniques and data from different market.
- Instead of feature selection of the time series data this model acts for denoising as well as feature extracting method.
- Although there are different DL techniques that has been previously imposed for the purpose of prediction, in this paper the DL is combined with the ML to accomplish more precise predicted result.
- Lastly DL has not been combined with the KELM technique for prediction in the field of banks or real estate hence this study makes an attempt to predict the financial market of different banks.

The main motive behind this study is to design a model that gives minimum prediction error and maximum profitability, different comparative study emphasizes that the proposed model outperforms various other methods in terms of prediction accuracy thereby leading to maximum profitability for investors.

Organization of the rest of the paper is given as follows: Section 2 shows the related work, mainly focusing on promising techniques proposed by different researchers for predicting financial market in order to obtain best result. Section 3 gives a brief description of the related study including the basic ELM, robust ELM and Kernel ELM. Section 4 introduces the prediction methodology implementing stacked AE in combination with the KELM classifier, followed by the data characterization and performance evaluation criteria in Section 5. Section 6 performs the analysis of the empirical results and Section 7 concludes the proposed study.

2. Related work

Although researchers have proposed different prediction models like ARMA, ARIMA, BP, FLANN, ELM, etc. although the traditional BP technique proved to be efficient for solving different complex problems but suffers from the problem of local minima and slower convergence speed hence the ELM model became more popular and significant for effective prediction of stock market. The main drawbacks of this model are selection of the hidden layer, and decrease in prediction accuracy when a larger data set with different input variables is taken into consideration. The limitation of hidden layer selection is vindicated by introducing the kernel matrix and transforming the basic ELM into Kernel based ELM making the day ahead or month ahead prediction more efficient. The KELM technique has shown better generalization performance for various real applications, apart from various advantages the major issue regarding the kernel is random choosing of various parameters of different kernel functions. The complications occurring because of the variation of input variables and larger data size is resolved up to a great extent with the use of auto encoder learning technique, where different layer reduces the size of the input matrix and thus providing a promising predicted output of the financial market. In this study the decision making layer comprises of the KELM method. . In ref [54] the authors uses the linear regression and SVM regression to predict the open price of the next day and closing price of the previous day for NSE stock market respectively. In ref [55] the use of FLANN and BP model is shown for predicting the stock market data of NSE, BSE and INFY. The New York stock exchange was predicted using the ARIMA method as shown in ref [56]. [57] Shows a study of predicting NSE data using different predicting method. A more recent study is performed using the DL technique for prediction of different NSE and NYSE financial market data, from NSE market TATA MOTORS is taken into consideration using MLP, RNN, LSTM and CNN methods. There are other financial market data from both NSE and NYSE those are considered for prediction and compared with the ARIMA model to observe the performance accuracy of the DL technique. Another recent study was performed where the authors integrated the algorithm with SAE and SVR for foreign exchange prediction and it was proved that the proposed model outperformed various other benchmark models [58].

3. Methodology

This section describes the basic of different AI models applied to perform the prediction and also for the purpose of comparison. ELM, robust KELM and deep learning technique is explained elaborately. This paper presents a DL framework for time series prediction of financial data, an integration of stacked AE and robust KELM technique. This structure is made up of three stages: (1) Data normalization (2) application of SAE and (3) decision making at the final stage with the help of RKELM. N Fig. 1(a) demonstrates the method framework and a flow chart describing the research methodology is shown in Fig. 1(b).

3.1. Kernel Extreme Learning Machine (KELM)

The basic ELM is a SLFN (single layer feed forward neural network) composed of three layers, the input layer the hidden layer and the output layer. The output function of the ELM with L hidden nodes is expressed as:

$$f_L(x) = \sum_{i=1}^{L} \beta_i h_i\,(x) = h\,(x)\,\beta \tag{1}$$

where $\beta = [\beta_1, \beta_2, \ldots, \beta_L]$, demonstrates the output vector forming a link between the hidden layer node (L) and output neuron, and

$h(x) = [h_1(x), h_2(x), \ldots, h_L(x)]$ Shows the ELM feature mapping function;

The number of input samples are denoted as $x = [x_1, x_2, \ldots, x_N]$, where N represents the number of patterns. As there is no need of tuning of a hidden layer, the parameters are randomly generated. The input data is further transformed to the hidden layer space with the help of feature mapping, an activation function maps the data, thus $h_i(x)$ can be written as

$$h_i(x) = a(w_{i0} + w_{i1}x_{i1} + w_{i2}x_{i2} + \cdots + w_{im}x_{im}) \tag{2}$$

Thus the hidden layer matrix can be formulated as

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} h(x_1) & \cdots & h_L(x_1) \\ \vdots & \vdots & \vdots \\ h(x_N) & \cdots & h_L(x_N) \end{bmatrix} \tag{3}$$

and the target vector can be written as

$$T = \begin{bmatrix} t_1 \\ t_2 \\ \cdot \\ \cdot \\ t_N \end{bmatrix} \tag{4}$$

Eq. (1) expressed in a matrix form can be written as

$$H\beta = T \tag{5}$$

The main motive of the ELM technique is to improve the performance generalization and reduce the norm of the output weight and training error in order to maintain the system stability. But the main drawback of the ELM technique being the choice of hidden layer that influences stability.

To solve for β, a constrained optimization problem needs to be solved that overcomes the over fitting problem and provides better generalization ability in comparison to the original ELM. This is similar to the structural risk minimization of the statistical learning theory and is expressed as

$$L_{C_{ELM}} = \frac{1}{2}\,\|\beta\|^2 + \frac{1}{2}C\,\|\xi\|^2 \tag{6}$$

subject to $H\beta = T - \xi$

where the error vector ξ is denoted as: $\xi = [\xi_1, \xi_2, \ldots, \xi_N]$, and C is the regularization parameter, selected between 2^5 to 2^{30} Using KKT theorem the constrained optimization problem given in Eq. (6) can be converted to dual optimization with the help of KKT, and can be written as

$$L_{D_{ELM}} = \frac{1}{2}\,\|\beta\|^2 + \frac{1}{2}C\,\|\xi\|^2 - \alpha(H\beta - T + \xi) \tag{7}$$

where $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_N]$ is the vector of Lagrange multipliers. Now solving Eq. (7) for optimality condition, the value of the vector β is obtained as

$$\beta = (\frac{I}{C} + H^T H)^{-1} H^T T \text{ for } N > L \tag{8}$$

$$\text{or } \beta = H^T(\frac{I}{C} + HH^T)^{-1} T \text{ for } N < L \tag{9}$$

where I is an unit matrix of appropriate dimension. The output for a sample x comprising m inputs from the ELM is written as

$$f(x) = h(x)(\frac{I}{C} + H^T H)^{-1} H^T T \text{ for } N > L \tag{10}$$

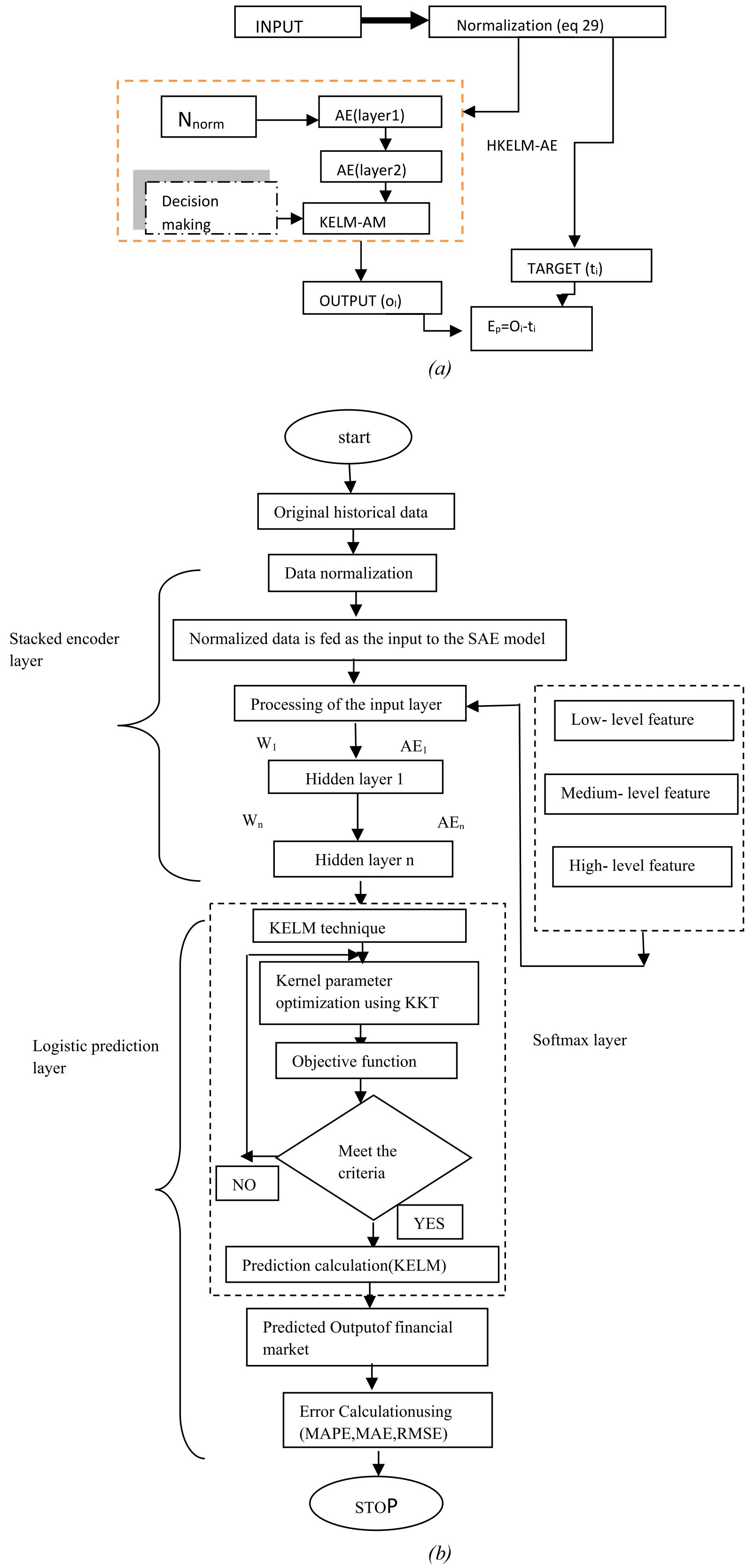

Fig. 1. Demonstration of the proposed model, (a). Structural demonstration of the proposed technique, (b) flow chart representing the methodology.

$$\text{or } f(x) = h(x)H^T(\frac{I}{C} + HH^T)^{-1}T \text{ for } N < L \tag{11}$$

Thus the total output vector for N number of patterns becomes

$$O_p = H(\frac{I}{C} + H^T H)^{-1} H^T T \text{ for } N > L$$

$$\text{And } O_p = HH^T(\frac{I}{C} + H^T H)^{-1} T \text{ for } N < L$$

In our case we consider $N > L$, the number of neurons in the hidden layer affects the learning performance, stability and generalization of the ELM, a proper selection of activation function which is still an unsolved problem. Thus forum known feature mapping function, Kernel functions can be used for the ELM providing better stability and generalization and the ELM with the newness can be designated as Kernel Extreme Learning Machine or simply KELM. Thus the Kernel theorem is given by

$$K_{ELM} = HH^T, \text{ and } K_{ELM}(x_i, x_j) = h(x_i)h(x_j) \tag{12}$$

Thus the output function is obtained as

$$f\,(x) = h\,(x)\,H^T\left(K_{ELM} + \frac{I}{C}\right)^{-1} T \tag{13}$$

Eq. (13) is written in an elaborated form as

$$f(x) = [(x, x_1), (x, x_2), \ldots, (x, x_N)](K_{ELM} + \frac{I}{C})^{-1} T \tag{14}$$

Kernel functions satisfy the Mercer condition and are convenient to use. Different Kernel functions that can be considered are Polynomial kernel; Gaussian Kernel, Hyperbolic tangent kernel (Sigmoid kernel) and Wavelet kernel. Different kernel functions considered are described as follows:

(1) Polynomial kernel: $K(x_i, x_j) = (1 + x_i^T x_j/\sigma^2)^g$ (15)

(2) Gaussian kernel: $K(x_i, x_j) = \exp(-\left\|(x_i - x_j)\right\|^2)/(2\sigma^2)$ (16)

(3) Sigmoid kernel: $K(x_i, x_j) = \tanh(bx_i^T x_j + c)$ (17)

(4) Wavelet kernel: $K(x_i, x_j) = \cos\left(\frac{d\,\left\|(x_i - x_j)\right\|}{e}\right) \times \exp\left(-\frac{\left\|(x_i - x_j)\right\|^2}{f}\right)$ (18)

The various parameters like σ, bc, d, e, f and γ are to be chosen appropriately to enhance the performance of the KELM based prediction model. By considering previous studies of the author the value of g is considered to be 2.

For the KELM with kernel function $K(x_i, x_j)$ the equation is written as following:

Further it is well known that for distinct N samples

$$\|K_{N\times N}\beta - T\|^2 = 0 \tag{19}$$

However, with L number of hidden layer samples

$$\|K_{N\times L}\beta - T\|^2 < \varepsilon \tag{20}$$

where ε is non zero training error and N is always greater than L. The error is minimized by using ℓ_2 norm minimization and sample number constraint, hence the optimization problem can be formulated as

$$\text{Minimize} \quad \frac{1}{2}\|\beta\|^2 + \frac{1}{2}C\,\|(K_{N\times L}\beta - T)\|^2 \tag{21}$$

Therefore the output vector can be written as

$$\beta = [\frac{I}{C} + K_{N\times L}^T K_{N\times L}]^{-1} K_{N\times L} T \tag{22}$$

Thus the predicted output is obtained as

$$f(x) = \begin{bmatrix} K(x, x_1) \\ K(x, x_2) \\ \cdot \\ \cdot \\ \cdot \\ K(x, x_L) \end{bmatrix}^T [\frac{I}{C} + K_{N\times L}^T K_{N\times L}]^{-1} K_{N\times L}^T T \tag{23}$$

3.2. Deep learning

This is a type of machine learning technique that possess the ability to extract the deep features mechanically. They mostly consist of more than 3 layers and abstract data or signal in a layer by layer manner and maintain a hierarchy of feature extraction that is from low level to high level [30]. As discussed earlier in Section 1 this is a more efficient machine learning technique with better feature extracting capability. The deep learning provides the output using the high level feature classification.

The different types DL can be sub divided as follows:

1. Deep Boltzmann Machine
2. Deep Belief Networks
3. Restricted Boltzmann Machine
4. Convolution Neural Network
5. Auto Encoder

In this paper we emphasis on the AE for the purpose of price prediction

3.3. 'Stacked auto encoder

As the name suggests in this kind several AE are stacked one above the other. It practically initializes the deep neural network. It is constructed by staking a number of single layers AE, the first layer is trained and the reconstruction layer is formed and further this reconstruction layer is fed as the input to the next layer. The hidden layer of the current AE acts as the input to the consequent AE. In this study AE is implemented for layer wise training of the input samples (OHLC) and the architecture of SAE is applied. It is a minimum three layer structure where the first layer represents the input layer and the last layer represents the reconstruction layer whereas the second layer represents the hidden layer. The main motive behind the training of every layer of AE is error reduction between the input and the reconstruction layer. It a two step process where the first step is to map the input layer and the hidden layers while the second step is to construct the reconstruction layer. The mathematical representation of the steps is shown below:

$$h(x) = a\,(W_1 x + b_1) \tag{24}$$

$$x' = a\,(W_2 h(x) + b_2) \tag{25}$$

where x and x′ represents the input and the reconstruction data respectively, W_1 and W_2 are the weight matrix of the hidden layer and reconstruction layer. b_1 and b_2 gives the bias values of the hidden layer and reconstruction layer respectively, a represents the activation function, although it can be of different alternatives in this paper we have considered the sigmoid in accordance to Ref [59]. The structural overview of a SAE is demonstrated as follows:

Step 1: the single layered AE is first mapped to the hidden layer and the reconstruction layer is obtained.

Step 2: Now the reconstruction layer is withdrawn and the hidden layer acts as the new input layer to the subsequent single layer AE. Same process is continued unless the desired output is obtained.

Most of the DL methods previously implemented used the BP model and thus the need of fine tuning of the hidden parameters

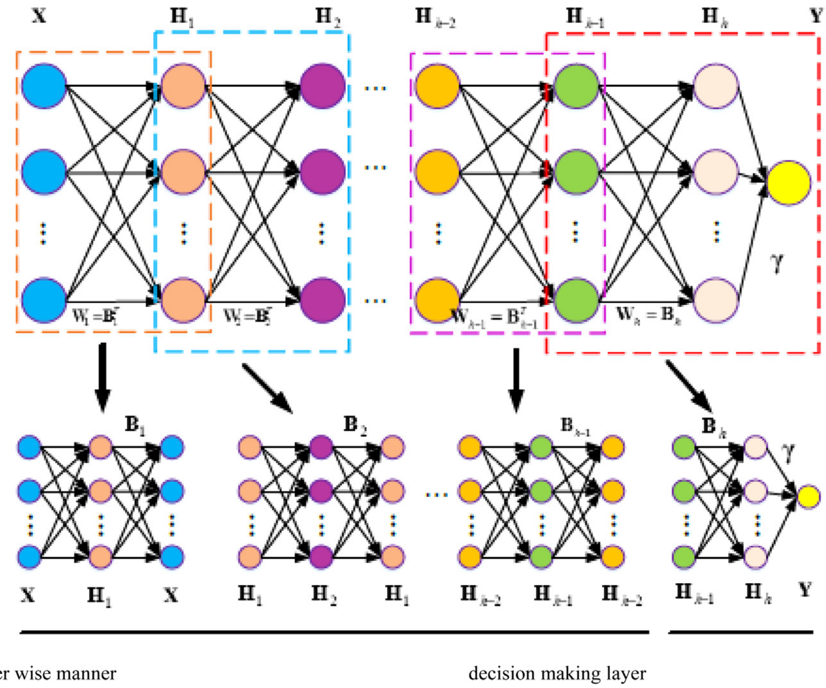

Fig. 2. Structural review of ELM-AE.

makes it a very much time consuming process hence the introduction of AE-ELM or hierarchical ELM, which performs efficiently in a much lesser time. Like the basic AE this consists of two steps where in the first step the AE-ELM is used to form the stacked architecture followed by the basic ELM in the last step for decision making purpose, in this method there is no need for parameter tuning and hence performs at a much higher speed as compared to the previous BP method.

3.4. Hierarchical ELM

It is a combination of two constituent, the ELM based feature encoding and classifier based on ELM for decision making purpose. The feature extraction is performed implementing the SAE process. HELM is a multi layer FFNN and the parameters are selected on the basis of multiple layers of the AE-ELM. It is a hybridization of the high learning efficiency of the ELM model and layer wise deep structure of the AE method, where the AE functions as the feature extractor of the multiple neural networks. Fig. 2 gives a structural review of the AE-ELM where it clearly shows the layer wise process performed by the AE-ELM, where the input samples are mapped to form the hidden layer which is further used to give the reconstruction layer and in the next step the reconstruction layer is removed and the hidden layer of the previous step acts as a input to the new hidden layer and the continuation of the process is performed until the minimized matrix is formed for the purpose of stock market prediction. HELM is a much faster performing technique as compared to the BP model. The result analysis section gives a proper view of the proposed model.

3.5. Multi-layer hierarchical KELM-AE

These functions similar to the HELM technique, but the only difference being the choice of the hidden layer where the randomized ELM-AE got replaced with a kernel matrix. The main advantage of the KELM technique is that there is no need for selection of hidden layer and hence more accurate prediction of the financial market can be observed. As a new contribution to the previously implemented technique the KELM-AE has not yet been included for the prediction of stock market. The K matrix replaces the ELM-AE classifier.

4. Data description

This section gives the detail knowledge regarding the selection of various input samples and the corresponding resources. This gives details of different market data from where the OHLC price is considered for the purpose of prediction of closing price.

4.1. Input variable selection

In this study five stock market indices are taken into consideration namely Yes Bank, SBI and bank of India Ltd, these data has been collected from the Indian Stock Market (NSE) [60]. Wider market gives a better opportunity demonstrating the capability of the proposed prediction technique. The Indian financial market is considered to be a new and developing market thus facing more data fluctuation and hence requires a better prediction technique. In addition to the above market a cross validation of the prediction model is performed using different stock market index. The stock market data from a developed market is chosen to prove the robustness of the proposed model. DowJones from the US Stock Market and ASHR is considered from Chinese Stock Exchange [61]. In accordance to the previous literature in this paper we have considered four different input variables namely the Open Price, High, Low &Closing Price of individual financial market for predicting the subsequent closing price. These variables demonstrate the market scenario of different stock market. An efficient stock market prediction helps the investors to understand the financial market in a more appropriate manner. 7 years data ranging from 6/11/2013 to 12/6/2020 is considered for calculating performance accuracy of the proposed KELM-AE method. We use 24 h high frequency data including the opening price, the high price, the low price and the closing price in order to investigate the performance of the proposed method and other AI techniques.

Algorithm of the proposed model

Start

Data =fmp (financial market price feed as input to the model)

N=length (fmp)

N_{norm}=fmp(i)- min(fmp)/max(fmp)-min(fmp)

Initially consider only 3 layers of SAE, i.e. the input layer x, the hidden layer h and output layer x', h_1 is the first hidden layer

Map the input layer with the first hidden layer

$h_1(x_1) = a(w_1 x_1 + b_1)$

W= randomly selected weight

X_1'= output layer obtained using eq 25

Output obtained is fed as the input to the next layer

Calculation of the second hidden layer

$h2(x2) = a(w_2 x2 + b2)$

$x_2 = x_1'$

Continue the process for another layer

The last layer/decision making layer is constructed with the output obtained from the last layer

Select the training and testing data

$trn = x_f(1:1163)$

$tst = x_f(1164:1662)$

for j=length (*tst*)

KELM model is implemented along with the selected kernel function as given in section 3.1 using eq 12, 13 and 14 with a particular type of kernel function given in eq 15/16/17 and 18

Select the kernel parameter and optimize using KKT

Start testing

for k= length (tst)

implement KELM for the total testing time period

end

Calculate PO_{test} (k)

Obtain AO(k)

Compare PO_{test} (k) & AO(k)

Error(k) = PO_{test} (k) - AO(k)

Calculate RMSE using eq. (28)

Calculate MSE using eq. (27)

Calculate MAPE using eq. (26)

Stop

A total number of 1662 samples with 4 variables create a matrix of 6648 data. Fig. 3 gives a graphical representation of 1 day ahead of the original historical data of YES bank. From the original signal, we can well understand the fluctuations present in the closing price data which in return leads to higher prediction error. Thus a precise prediction model needs to be proposed for a more accurate prediction of the financial market and the stepping stone to this being data normalization for maintaining the uniformity of the prediction data.

Performance Evaluation

Various performance indices are considered for testing the performance accuracy of the proposed model and other intelligence technique, the lesser the error the better is the prediction

Fig. 3. Original 24 h based closing price data for YES bank.

accuracy. Some of the performance measuring units are given below:

- Mean Absolute Percentage Error (MAPE)

$$MAPE = \left(\frac{1}{N}\sum\sum_{i=1}^{N}\frac{|t_i - o_i|}{t_i}\right) * 100 \tag{26}$$

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|t_i - o_i| \tag{27}$$

- Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_i - o_i)^2} \tag{28}$$

$$N_{norm} = \frac{N - N_{\min}}{N_{\max} - N_{\min}} \tag{29}$$

where the normalized data is denoted as N_{norm}, N being the total input data, N_{min} and N_{max} are the minimum and maximum price of the entire data set considered for prediction, t_i is the target and O_i is the measured output.

5. Empirical analysis of result

In this paper 7 years data of various financial market belonging to different region has been taken into consideration for the purpose of prediction of the financial market. The whole data is initially normalized between 0 to 1 using Eq. (29), the normalized data rules out the concept of data overlapping, and ranging of the data within a certain limit maintains the uniformity of the prediction irrespective of the data location. The normalized data is further divided into two groups training and testing with 7:3 ratio, i.e. 70% of the total data set is considered for training purpose whereas the rest 30% is considered for testing. The experimental approach is conducted in MATLAB 15 software in a system possessing the specification i5, 64 bit, 4 GB ram and CPU 7200U @2.50 GHz.

DL techniques generally consists of more than three layers where the initialization is performed with the help of unsupervised learning method and the tuned using supervised training that progressively generates more high level features in a layer by layer manner. In this study layer wise training of the OHLC variables are performed. The main aim of training the single layer AE

Table 1
MAPE (%) value for different kernel function for different year.

Year	Polynomial	Gaussian	Wavelet	Sigmoid
2013	1.22	2.15	1.38	1.72
2014	1.31	2.14	1.34	1.77
2015	1.35	2.27	1.43	1.82
2016	1.32	2.24	1.41	1.85
2017	1.34	2.16	1.39	1.88

Table 2
Comparative study of different prediction technique for price prediction of YES bank.

Techniques	MAPE (%)	MAE	RMSE	R	T_t (s)	C_t (s)
DL (KELM-AE)	**1.074**	**0.010**	**0.016**	**0.998**	**8.383**	**24.293**
ELM	1.714	0.017	0.025	0.996	1.085	7.462
RBF	3.745	0.037	0.059	0.981	6.442	18.873
BP	4.424	0.044	0.070	0.956	5.231	20.871

C_t is the computational time and T_t is the training time both given in sec.

Table 3
Comparative study of different prediction technique for price prediction of SBI bank.

Techniques	MAPE (%)	MAE	RMSE	R	T_t (s)	C_t (s)
DL (KELM-AE)	**0.888**	**0.009**	**0.023**	**0.998**	**7.03**	**23.940**
ELM	1.473	0.017	0.025	0.996	1.82	8.054
RBF	2.412	0.024	0.051	0.991	7.11	10.203
BP	3.962	0.039	0.074	0.996	7.90	25.677

C_t is the computational time and T_t is the training time both given in sec.

Table 4
Comparative study of different prediction technique for price prediction of BOI bank.

Techniques	MAPE (%)	MAE	RMSE	R	T_t (s)	C_t (s)
DL (KELM-AE)	**1.021**	**0.0163**	**0.023**	**0.998**	**5.935**	**23.253**
ELM	1.550	0.015	0.021	0.997	0.462	7.869
RBF	2.383	0.034	0.031	0.994	6.982	14.578
BP	3.228	0.032	0.035	0.996	7.440	25.037

C_t is the computational time and T_t is the training time both given in sec.

is error reduction between the input layer and the reconstruction layer. In this study KELM technique is used as the final layer or the decision making layer, where there is no need for the formation of the H matrix as in case of the conventional ELM because the H matrix is now replaced with a kernel matrix. Different types of kernel functions can be combined with the basic ELM method but in this case the Polynomial kernel is considered as the performing kernel function. Table 1 gives a comparative study of different kernel function based AE for YES bank's price prediction for a time interval of 1 year and it can be observed that the polynomial kernel function gave better prediction accuracy based on MAPE value as compared to the other functions hence considered as the performing kernel function for rest of the case studies performed in this paper. Table 2 gives a comparative study of different AI technique along with the proposed DL (KELM-AE) technique for predicting the financial market of YES bank. Different performance matrices along with computational time for training the data and running the whole algorithm is properly mentioned here, based on different evaluation criteria it can be clearly noted that the proposed method outperformed other conventional methods like ELM, RBFNN and BP. The corresponding graphical validation is given Fig. 4, over here we can clearly observe the variation in prediction accuracy, and more the distance of the actual from predicted signal using different prediction method less accurate is the prediction model. The original market variation of YES bank can be observed from Fig. 3 where the closing price corresponding to each day is illustrated.

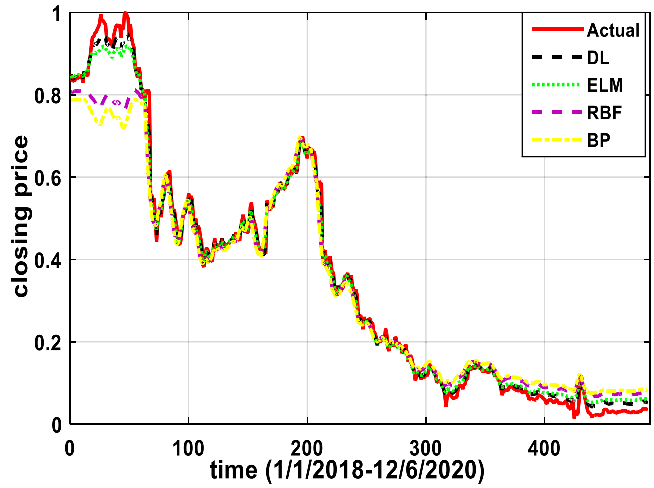

Fig. 4. Prediction comparison: Price prediction of YES using different prediction models.

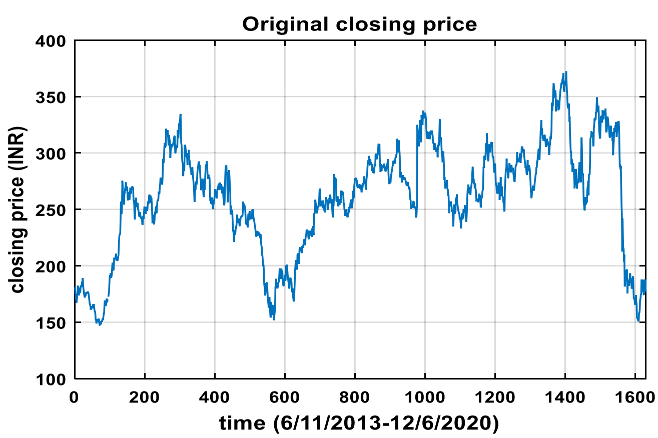

Fig. 5. Original 24 h based closing price for SBI bank.

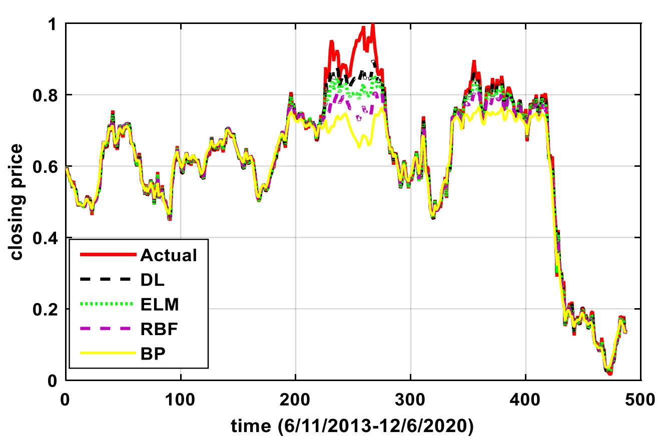

Fig. 6. Prediction comparison: Price prediction of SBI using different prediction models.

In a similar manner Table 3 demonstrates a comparative study of the proposed DL technique (KELM-AE) with other mentioned techniques for predicting 7 years data of SBI. Original closing price of SBI and the comparison graph emphasizing the prediction accuracy of different techniques along with the proposed model is shown in Figs. 5 and 6 respectively, which shows a similar tendency in prediction like the YES bank data prediction. Another

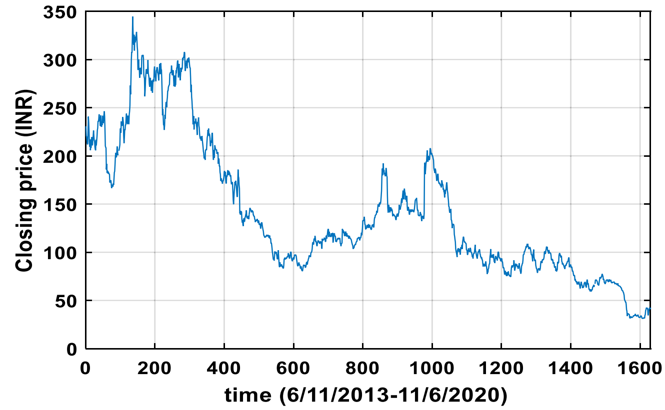

Fig. 7. Original 24 h based closing price for bank of India.

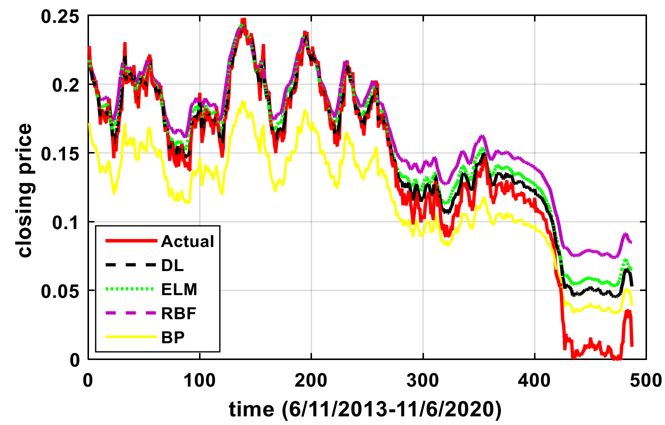

Fig. 8. Prediction comparison: Price prediction of BOI using different prediction models.

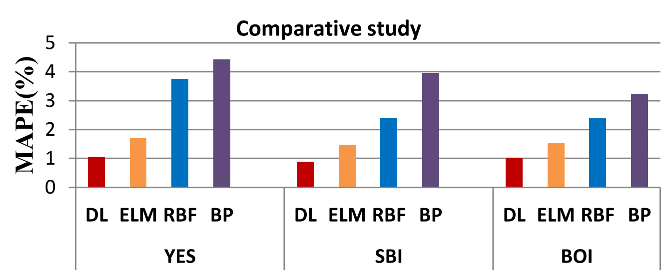

Fig. 9. Comparative study of MAPE for different prediction model for different market.

experiment is performed for Bank of India, the original closing price data for Bank of India is shown in Fig. 7 and the comparative study between different prediction methods is given in Fig. 8 and the corresponding table showing different values of the performance indices is described in Table 4. A pictorial representation of MAPE values of the proposed DL baring to different market data is shown in Fig. 9 where the robustness of the model is well observed, i.e. irrespective of the financial market the proposed model provides a great accuracy in prediction. The KELM-AE (DL technique) gives better result as compared to other prediction models compared on the basis of its MAPE percentage, the MAPE value of the predicted technique is minimum irrespective of the financial market.

Two more case studies are performed for validating the proposed model, where two different financial market one being the Chinese market (ASHR) and the other being the NY stock

Fig. 10. Original 24 h based closing price for ASHR.

Table 5
Comparative study of different prediction technique for price prediction of ASHR — Chinese stockmarket data.

Techniques	MAPE (%)	MAE	RMSE	R	T_t (s)	C_t (s)
DL (KELM-AE)	**0.696**	**0.007**	**0.009**	**0.998**	**5.964**	**27.592**
ELM	1.013	0.010	0.0133	0.997	1.025	8.735
RBF	3.160	0.032	0.037	0.967	5.384	16.772
BP	3.617	0.036	0.042	0.977	8.838	25.612

C_t is the computational time and T_t is the training time both given in sec.

exchange (DJI), considered to be a developing and developed market respectively, are predicted. Total number of input data for the ASHR and DJI market being 1662 with four different variable set as input. Original historical closing price of the above mentioned data is shown in Figs. 10 and 11 respectively. In order to authenticate the proposed technique it was implemented to predict the stock market of two different financial market out of which ASHR belongs to a developing market and DJI is a developed market Tables 5 and 6 shows a comparative study of different prediction technique when applied to the respective financial markets and it can be well observed that the DL(KELM-AE) method outperforms all other prediction models irrespective of the financial data thus verifying the robustness of the proposed model, the best result is obtained using the KELM-AE followed by the ELM, BP and so on. The corresponding graph is shown in Figs. 12 and 13, where a similar prediction trend to that of YES bank, SBI and BOI is observed thus maintaining the robustness of the model. The developed market data possesses greater accuracy as compared to the developing markets. In addition to this another case study has been performed where a larger number of input data is considered and it was obtained that the DL method functioned well for larger data set. In accordance with ref [49] where the authors have mentioned a wide literature review of various previously implemented models, in this paper some of the techniques mentioned are compared with the proposed model on the basis of statistical test, for the DJI stock market as it is considered to be a developed market. The main complexity while using this technique is the choice of the kernel function and assigning the parameter values for obtaining the optimal result, this should be letter solved in future work by introducing optimization technique for parameter optimization.

5.1. Profitability analysis

A profitability performance is performed to find the most valuable model earning the highest profit for the investors and in order to do so the buy and sell trading is performed. The profitability is checked by comparing the prices of the present

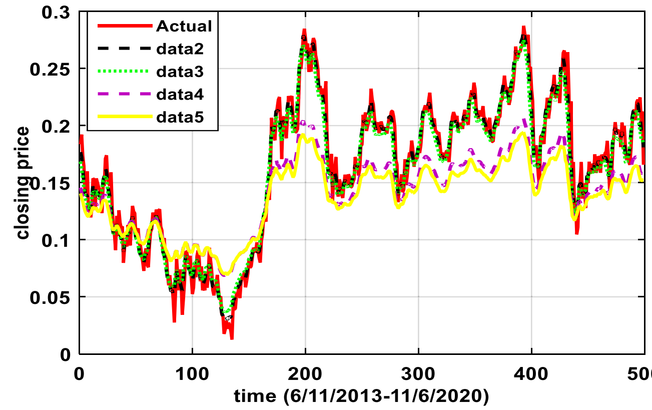

Fig. 11. Prediction comparison: Price prediction of ASHR market using different prediction models.

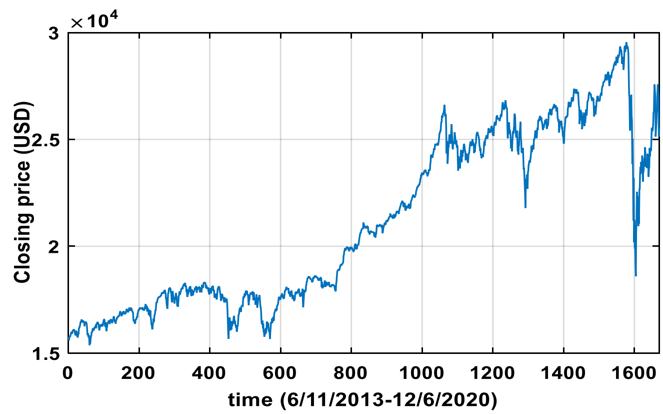

Fig. 12. Original 24 h based closing price for DJI market.

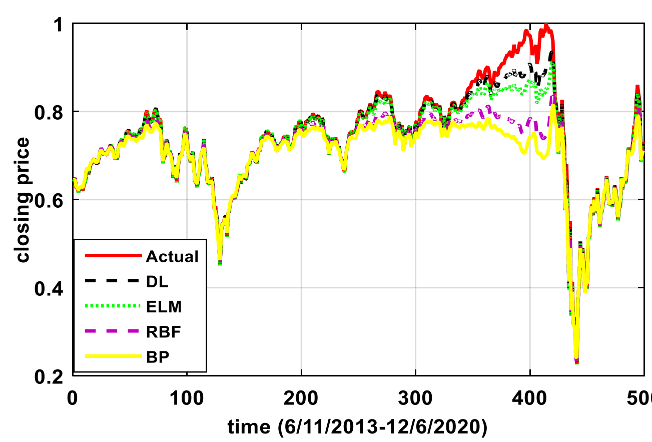

Fig. 13. Prediction comparison: Price prediction of DJI market using different prediction models.

Table 6
Comparative study of different prediction technique for price prediction of DJI market — NYSE stock market data.

Techniques	MAPE (%)	MAE	RMSE	R	T_t (s)	C_t (s)
DL (KELM-AE)	**0.848**	**0.008**	**0.021**	**0.997**	**4.347**	**26.476**
ELM	1.396	0.014	0.032	0.989	0.866	8.429
RBF	3.096	0.031	0.062	0.977	6.297	18.813
BP	4.109	0.041	0.077	0.963	8.225	22.381

C_t is the computational time and T_t is the training time both given in sec.

Table 7
Comparative study of four different prediction model based on statistical tests.

Prediction model	Mean rank (Friedman test)	Mean rank difference	p-value
KELM-AE	3.0808	–	–
ELM	3.9897	−0.9089	0.0000082
SVR	4.4270	−1.3462	0.0000001
RBF	4.8423	−1.7615	0.0000000
BP	8.1172	−5.0364	0.0000000

The p-values for the Dunn–Sidak test are shown for a level of seven decimal points and maximum of them consists of all zeros while others are shown with very small values less than 0.05 which proves that the performance of the proposed KELM-AE model is significantly different to all other predictive models considered in this study.

and following days, i.e. we can predict the buying and selling capability of the investors and hence loss minimization. The buy and sell strategy can be performed with the help of the following equation. The buying and selling depends upon the following factor

$$\begin{aligned} &d_{t+1} > d_t \rightarrow buy \\ &d_{t+1} < dt \rightarrow sell \end{aligned} \tag{30}$$

Where d_t denotes the present actual value and d_{t+1} is the predicted value of the next time period. And the strategy earnings can be defined as

$$\begin{aligned} R_s = 100 \times \Bigg(&\sum_{t-1}^{b} \frac{d_{t+1} - d_t + (d_t * B_c + d_{t+1} * S_c)}{d_t} \\ &+ \sum_{t-1}^{r} \frac{d_t - d_{t+1} + (d_{t+1} * B_c + d_t * S_c)}{d_t} \Bigg) \end{aligned} \tag{31}$$

Here R_s denotes the strategy returns, s and b is the total number of days for selling and buying respectively and S_c and B_c represents the transaction cost for selling and buying respectively. The trading strategy is performed based on the standard future products. The profitability of DJI market data is shown in Table 8, the result suggests that the KELM-AE substantially earns more profit than the other models. Thus the superiority of the proposed model is confirmed even in terms of profitability test.

5.2. Result outcome discussion

Five data set such as YES bank, SBI bank, BOI bank of Indian stock market, ASHR of China stock market and DJI of New York stock market are utilized for observing the ability of the proposed prediction model, this individual data set is divided into training and testing set. The main aim of the kernel based ELM accompanied with the deep learning technique is to find a suitable solution with a considerably less run time. The SAE is a unsupervised learning technique that learns the financial time series features. In this technique multiple single layered AE are used where the output features wires to the input feature of the next layer. The unsupervised training of the SAE reduces the error between the output data and the successive input data, thereby resulting in invariant learning and abstracting features. Data from three different markets are considered for obtaining the robustness of the model. In contradiction to the EMH theory that states that the predictability of a market asset is affected by its efficiency. This study helps in achieving this goal where two stages of development of financial markets are considered, the stock market of India and China is considered to be the developing market in contrast to that, the New York stock market is recognized to be the developed market and by far the largest stock market, thus the validity of our proposed method is validated in different market states. Following the previous literature study four historical trading data is considered for the purpose of predicting the closing price. The profitability result suggests that the proposed KELM-AE substantially earns more profit as compared to other methods. The experimental outcome demonstrated in the figure shows the performance accuracy of the proposed technique, as discussed earlier the KELM method used for the decision making purpose performs better than the traditional ELM as the problem of choosing the hidden layer is eliminated, the comparative graph shows the difference between the actual signal and the predicted signals using various other techniques. Different performance matrices confirms the performance efficiency of the proposed technique for example the Coefficient Correlation ® mentioned in the comparison table proves the closeness of the proposed signal with the actual signal, as we know closer the value towards 1 more accurate is the technique, similar trend is followed by other performance matrices that shows lesser error.

5.3. Prediction model comparison using nonparametric test

As per various previous studies in this paper two different nonparametric tests namely the Dunn–Sidak test and Friedman test [62,63] are performed to obtain the prediction accuracy of the proposed model. The test is conducted with 95% confidence level for testing the null hypothesis. The DJI market is considered for performing the test. The test statistics is observed to be quite less and the p value is obtained to be 1.0512e−278 which signifies the fact that the null hypothesis is rejected at a significant level of 5% and that a much significant difference can be observed in the performance accuracy of different model. Table 7 shows the mean ranking values and the p-values obtained from the Friedman test and Dunn–Sidak post-hoc procedure, on the basis of the rank and p value the proposed model proves to be the best performing statistical model as compared to other models, the proposed model show the p value less than 0.05 which proves the significant difference between KELM-AE model and other prediction models. Thus it emphasis that the proposed KELM-AE is better performing prediction model in case of stock market prediction.

6. Remarks

Interpretation of the financial market is of paramount importance, where different methods has been already proposed by various researchers with an expectation to meet the actual data. The empirical result shows the prediction of different financial markets from different location, the banks are considered from the NSE market which is one of the leading stock market and comprises of over 50 companies. Out of the vast category historical data of three banks are considered, for predicting the closing price using different methods, the minimum error is obtained using the DL method, i.e. the KELM-AE technique that performs with a minimum MAPE error of 1.074%, 0.888% and 1.02% for predicting the data of YES bank, SBI bank and BOI bank respectively, and the maximum MAPE value is observed using the BP technique where the MAPE value obtained is 4.424%, 3.962% and 3.228% for YES bank, SBI bank and BOI bank respectively. The same techniques were also conducted for studying two non-Indian financial market such as ASHR belonging to the Shanghai stock market and DJI belonging to the NYSE stock market and both showed similar trend while performing the prediction with the best performing KELM-AE with a MAPE value of 0.698% and 0.848% respectively whereas the maximum error is obtained using the BP model with a MAPE value of 3.617% and 4.109% respectively for the

Table 8
Profitability analysis of different prediction technique for DJI market.

Technique	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6	Average
KELM-AE	77.643	56.544	42.810	48.124	83.602	81.533	**65.043**
ELM	48.282	36.176	31.792	30.542	31.511	55.985	39.048
RBF	34.977	23.062	18.919	−4.953	5.310	38.689	19.334
BP	5.875	7.852	−2.083	9.640	38.457	18.912	13.109

above mentioned stock markets. From Table 8 it can be concluded that the proposed model shows a gain of almost 40% above the other models, for example when DJI market is considered it is observed that the annual earnings of the proposed model can reach up to an amount of 65.043% which almost 40% above the other three models. The computational time of the algorithm is computed using the tic toc command, it is implemented twice in the program where once it is used for calculating the training time of the individual prediction model and the second time it is used to calculate the total run time. The run time may vary for different run hence a total number of 10 runs are performed and the average run time is considered as the training time and the computational time.

7. Conclusion

Stock market prediction has proved to be challenging in nature because of the nonlinearity present in the input data, thus requires an accurate prediction model for predicting the future stock market. DL based method is well suited for this purpose as there are different input variables that influence the closing price, the deep learning technique shows better performance when different varieties of variables are given as input. In order to achieve a better stock market prediction performance we have used the deep learning architecture in combination with the KELM in its last layer for decision making purpose, the main drawback of the previously used ELM technique is the choice of hidden layer which is mitigated in this study by introducing the kernel function in the ELM that omits the difficulty in selecting the hidden layer network. In this paper we have compared the proposed method with various other previously used prediction methods like the ELM, BP, and RBF in accordance with different financial market of both developing and developed markets. The effectiveness of the proposed KELM-AE is verified by performing experiment on a total 7 years data for six different financial markets. Comparison with different previously implemented prediction model shows better efficiency of the proposed technique in terms of various performance matrices. In addition to this a data set with larger number of inputs (2005–2020) is considered for prediction, the selling and buying criteria were also looked upon to reduce the loss percentage of investors, the year wise buying and selling gives the profitability of the market. The MAPE value of YES bank using KELM-AE is obtained to be 1.05%, which is much less as compared to 1.714%, 3.745%, 4.424% and 2.53% obtained by using ELM, RBF, BP and SVM respectively. A similar trend is observed for different Indian stock market data using KELM-AE and MAPE value is observed to be 0.887% and 1.021% for SBI and bank of India ltd respectively, the MAPE values for DJI (NYSE) and ASHR (Chinese stock market) are observed to be 0.69% and 0.85% respectively, which is much less when compared to other prediction model. In future a more detailed study of the financial market will be performed where sudden drop in the market will be considered to authenticate the performance of the modified model considering the crash in the market due to various situation. And different meta heuristic optimization techniques will be added for better performance of the kernel parameters, where different case studies should be performed to demonstrate the robustness of the technique even in adverse situation. Table 9 shows a comparison of different similar type of study but applied in other field of study.

Table 9
Comparative study of different work performed.

Technique	MAPE	Reference
Proposed (DJI) (financial market prediction)	0.848	–
EVWCA-MKRVFLN (solar prediction)	1.06	[64]
Extreme SAE (energy consumption prediction)	3.0082	[65]
BPNN (energy consumption prediction)	4.1792	[65]
SAE_BP (electricity load prediction)	0.0488	[66]

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M.M. Rounaghi, F.N. Zadeh, Investigation of market efficiency and financial stability between S & P 500 and London stock exchange: monthly and yearly forecasting of time series stock returns using ARMA model, Physica A 456 (2016) 10–21.
[2] J.H. Wang, J.Y. Leu, June. Stock market trend prediction using ARIMA-based neural networks, in: Proceedings of International Conference on Neural Networks (ICNN'96), Vol. 4, IEEE, 1996, pp. 2160–2165.
[3] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, J. Econometrics 31 (3) (1986) 307–327.
[4] G.P. Girish, Spot electricity price forecasting in Indian electricity market using autoregressive-GARCH models, Energy Strategy Reviews 11 (2016) 52–57.
[5] L.A. Laboissiere, R.A. Fernandes, G.G. Lage, Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks, Appl. Soft Comput. 35 (2015) 66–74.
[6] A. Kavousi-Fard, A. Khosravi, S. Nahavandi, A new fuzzy-based combined prediction interval for wind power forecasting, IEEE Trans. Power Syst. 31 (1) (2015) 18–26.
[7] S. Karasu, A. Altan, Recognition model for solar radiation time series based on random forest with feature selection approach, in: 2019 11th International Conference on Electrical and Electronics Engineering (ELECO), IEEE, 2019, pp. 8–11.
[8] S. Karasu, A. Altan, Z. Saraç, R. Hacıoğlu, Estimation of wind speed by using regression learners with different filtering methods, in: 1st International Conference on Energy Systems Engineering, Karabuk, Turkey, 2017.
[9] M.P. Mohammed, River flood hazard modeling: forecasting flood hazard for disaster risk reduction planning, Civil Eng. J. 5 (11) (2019) 2309–2317.
[10] S. Ghorbani, M. Barari, M. Hosseini, A modern method to improve of detecting and categorizing mechanism for micro seismic events data using boost learning system, Civil Eng. J. 3 (9) (2017) 715–726.
[11] V. Mehdipour, M. Memarianfard, Application of support vector machine and gene expression programming on tropospheric ozone prognosticating for tehran metropolitan, Civ. Eng. J. 3 (8) (2017) 557–567.
[12] A.M. Rather, A. Agarwal, V.N. Sastry, Recurrent neural network and a hybrid model for prediction of stock returns, Expert Syst. Appl. 42 (6) (2015) 3234–3241.
[13] A.H. Moghaddam, M.H. Moghaddam, M. Esfandyari, Stock market index prediction using artificial neural network, J. Econ. Finance Adm. Sci. 21 (41) (2016) 89–93.
[14] S. Galeshchuk, Neural networks performance in exchange rate prediction, Neurocomputing 172 (2016) 446–452.
[15] I.P. Panapakidis, A.S. Dagoumas, Day-ahead electricity price forecasting via the application of artificial neural network based models, Appl. Energy 172 (2016) 132–151.
[16] C. Ren, N. An, J. Wang, L. Li, B. Hu, D. Shang, Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting, Knowl.-Based Syst. 56 (2014) 226–239.
[17] A. Karimzadeh, O. Shoghli, Predictive analytics for roadway maintenance: A review of current models, challenges, and opportunities, Civil Eng. J. 6 (3) (2020) 602–625.

[18] P. Srisuksomwong, J. Pekkoh, Artificial neural network model to prediction of eutrophication and microcystis aeruginosa bloom, Emerg. Sci. J. 4 (2) (2020) 129–135.
[19] M.Y. Chen, B.T. Chen, A hybrid fuzzy time series model based on granular computing for stock price forecasting, Inform. Sci. 294 (2015) 227–241.
[20] X. Ma, Y. Jin, Q. Dong, A generalized dynamic fuzzy neural network based on singular spectrum analysis optimized by brain storm optimization for short-term wind speed forecasting, Appl. Soft Comput. 54 (2017) 296–312.
[21] F. Ye, L. Zhang, D. Zhang, H. Fujita, Z. Gong, A novel forecasting method based on multi-order fuzzy time series and technical analysis, Inform. Sci. 367 (2016) 41–57.
[22] O.B. Shukur, M.H. Lee, Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA, Renew. Energy 76 (2015) 637–647.
[23] W. Kristjanpoller, M.C. Minutolo, Gold price volatility: A forecasting approach using the artificial neural network–GARCH model, Expert Syst. Appl. 42 (20) (2015) 7245–7251.
[24] D.C. Park, M.A. El-Sharkawi, R.J. Marks, L.E. Atlas, M.J. Damborg, Electric load forecasting using an artificial neural network, IEEE Tans. Power Syst. 6 (2) (1991) 442–449.
[25] J.L. Ticknor, A Bayesian regularized artificial neural network for stock market forecasting, Expert Syst. Appl. 40 (14) (2013) 5501–5506.
[26] L. Wang, Y. Zeng, T. Chen, Back propagation neural network with adaptive differential evolution algorithm for time series forecasting, Expert Syst. Appl. 42 (2) (2015) 855–863.
[27] E. Guresen, G. Kayakutlu, T.U. Daim, Using artificial neural network models in stock market index prediction, Expert Syst. Appl. 38 (8) (2011) 10389–10397.
[28] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: a new learning scheme of feedforward neural networks, in: 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541), Vol. 2, IEEE, 2004, pp. 985–990.
[29] W. He, T. Meng, S. Zhang, J.K. Liu, G. Li, C. Sun, Dual-loop adaptive iterative learning control for a timoshenko beam with output constraint and input backlash, IEEE Trans. Syst. Man Cybern.: Syst. 49 (5) (2017) 1027–1038.
[30] Y. Liu, H. Lu, K. Yan, H. Xia, C. An, Applying cost-sensitive extreme learning machine and dissimilarity integration to gene expression data classification, in: Computational Intelligence and Neuroscience, 2016, 2016.
[31] A. Iosifidis, A. Tefas, I. Pitas, Dropelm: Fast neural network regularization with dropout and dropconnect, Neurocomputing 162 (2015) 57–66.
[32] S. Shamshirband, K. Mohammadi, H.L. Chen, G.N. Samy, D. Petković, C. Ma, Daily global solar radiation prediction from air temperatures using kernel extreme learning machine: A case study for Iran, J. Atmos. Sol.-Terr. Phys. 134 (2015) 109–117.
[33] H. Fu, C.M. Vong, P.K. Wong, Z. Yang, Fast detection of impact location using kernel extreme learning machine, Neural Comput. Appl. 27 (1) (2016) 121–130.
[34] M. Pal, A.E. Maxwell, T.A. Warner, Kernel-based extreme learning machine for remote-sensing image classification, Remote Sens. Lett. 4 (9) (2013) 853–862.
[35] A. Altan, S. Karasu, The effect of kernel values in support vector machine to forecasting performance of financial time series, J. Cogn. Syst. 4 (1) (2019) 17–21.
[36] L. Deng, D. Yu, Deep learning: methods and applications, Found. Trends Signal Process. 7 (3–4) (2014) 197–387.
[37] X.W. Chen, X. Lin, Big data deep learning: challenges and perspectives, IEEE Access 2 (2014) 514–525.
[38] Y. Bengio, P. Lamblin, D. Popovici, H. Larochelle, Greedy layer-wise training of deep networks, Adv. Neural Inf. Process. Syst. (2007) 153–160.
[39] R. Katuwal, P.N. Suganthan, Stacked autoencoder based deep random vector functional link neural network for classification, Appl. Soft Comput. 85 (2019) 105854.
[40] X. Ding, Y. Zhang, T. Liu, J. Duan, Deep learning for event-driven stock prediction, in: Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015.
[41] L. Troiano, E.M. Villa, V. Loia, Replicating a trading strategy by means of LSTM for financial industry applications, IEEE Trans. Ind. Inf. 14 (7) (2018) 3226–3234.
[42] Z. Gao, W. Dang, C. Mu, Y. Yang, S. Li, C. Grebogi, A novel multiplex network-based sensor information fusion model and its application to industrial multiphase flow system, IEEE Trans. Ind. Inf. 14 (9) (2017) 3982–3988.
[43] M. Wang, L. Tian, From time series to complex networks: The phase space coarse graining, Physica A 461 (2016) 456–468.
[44] Z.K. Gao, Q. Cai, Y.X. Yang, N. Dong, S.S. Zhang, Visibility graph from adaptive optimal kernel time-frequency representation for classification of epileptiform EEG, Int. J. Neural Syst. 27 (04) (2017) 1750005.
[45] Z.K. Gao, S. Li, W.D. Dang, Y.X. Yang, Y. Do, C. Grebogi, Wavelet multiresolution complex network for analyzing multivariate nonlinear time series, Int. J. Bifurcation Chaos 27 (08) (2017) 1750123.
[46] G. Wang, J.R. Moffitt, X. Zhuang, Multiplexed imaging of high-density libraries of RNAs with MERFISH and expansion microscopy, Sci. Rep. 8 (1) (2018) 1–13.
[47] I. Sutskever, G.E. Hinton, Deep, narrow sigmoid belief networks are universal approximators, Neural Comput. 20 (11) (2008) 2629–2636.
[48] N. Le Roux, Y. Bengio, Deep belief networks are compact universal approximators, Neural Comput. 22 (8) (2010) 2192–2207.
[49] A.M. Ozbayoglu, M.U. Gudelek, O.B. Sezer, Deep learning for financial applications: A survey, Appl. Soft Comput. (2020) 106384.
[50] P.C. Chang, Y.W. Wang, W.N. Yang, An investigation of the hybrid forecasting models for stock price variation in Taiwan, J. Chin. Inst. Ind. Eng. 21 (4) (2004) 358–368.
[51] K. Parasuraman, A. Elshorbagy, Wavelet networks: an alternative to classical neural networks, in: Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005. Vol. 5, IEEE, 2005, pp. 2674–2679.
[52] A. Altan, S. Karasu, Recognition of COVID-19 disease from X-ray images by hybrid model consisting of 2D curvelet transform, chaotic salp swarm algorithm and deep learning technique, Chaos Solitons Fractals (2020) 110071.
[53] A. Altan, S. Karasu, The effect of kernel values in support vector machine to forecasting performance of financial time series, J. Cogn. Syst. 4 (1) (2019) 17–21.
[54] M. Gurjar, P. Naik, G. Mujumdar, T. Vaidya, Stock market prediction using ANN, Int. Res, J. Eng. Technol. (IRJET) 5 (03) (2018).
[55] P. Mohapatra, A. Raj, T.K. Patra, Indian stock market prediction using differential evolutionary neural network model, Int. J. Electron. Commun. Comput. Technol. (IJECCT) 2 (4) (2012) 159–166.
[56] A.A. Ariyo, A.O. Adewumi, C.K. Ayo, Stock price prediction using the ARIMA model, in: 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, IEEE, 2014, pp. 106–112.
[57] K.K. Sureshkumar, N.M. Elango, An efficient approach to forecast Indian stock market price and their performance analysis, Int. J. Comput. Appl. 34 (5) (2011) 44–49.
[58] M. Hiransha, E.A. Gopalakrishnan, V.K. Menon, K.P. Soman, NSE stock market prediction using deep-learning models, Procedia Comput. Sci. 132 (2018) 1351–1362.
[59] Y. Chen, Z. Lin, X. Zhao, G. Wang, Y. Gu, Deep learning-based classification of hyperspectral data, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (6) (2014) 2094–2107.
[60] http://www.nseindia.com.
[61] https://finance.yahoo.com/.
[62] M. Friedman, The use of ranks to avoid the assumption of normality implicit in the analysis of variance, J. Amer. Statist. Assoc. 32 (200) (1937) 675–701.
[63] B.E. Madsen, S.R. Browning, A groupwise association test for rare mutations using a weighted sum statistic, PLoS Genet. 5 (2) (2009) e1000384.
[64] I. Majumder, P.K. Dash, R. Bisoi, Short-term solar power prediction using multi-kernel-based random vector functional link with water cycle algorithm-based parameter optimization, Neural Comput. Appl. (2019) 1–19.
[65] C. Li, Z. Ding, D. Zhao, J. Yi, G. Zhang, Building energy consumption prediction: An extreme deep learning approach, Energies 10 (10) (2017) 1525.
[66] Y. Zou, M. Tu, X. Teng, R. Cao, W. Xie, Electricity price forecast based on stacked autoencoder in spot market environment, in: 2019 9th International Conference on Power and Energy Systems (ICPES), IEEE, 2019, pp. 1–6.